THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 9, pp. 6199–6211, February 28, 2014

A Novel Unsaturated β-Glucuronyl Hydrolase Involved in Ulvan Degradation Unveils the Versatility of Stereochemistry Requirements in Family GH105*

Received for publication, December 6, 2013, and in revised form, January 9, 2014 Published, JBC Papers in Press, January 9, 2014, DOI 10.1074/jbc.M113.537480

Pi Nyvall Collén[‡1], Alexandra Jeudy[‡], Jean-François Sassi[§], Agnès Groisillier[‡], Mirjam Czjzek[‡], Pedro M. Coutinho[¶], and William Helbert[‡‖2]

From the [‡]CNRS, Université Pierre et Marie Curie-Paris 6, Unité Mixte de Recherche 7139 "Marine Plants and Biomolecules," Station Biologique, F-29682 Roscoff Cedex, France, [§]Centre d'Etudes et de Valorisation des Algues, Presqu'île de Pen Lan, BP3, 22610 Pleubian, France, [¶]Architecture et Fonction des Macromolécules Biologiques, Unité Mixte de Recherche 7257, CNRS, Aix-Marseille Université, 163 Avenue de Luminy, Marseille 13288, France, and [‖]Centre de Recherches sur les Macromolécules Végétales (UPR-CNRS 5301), Université Joseph Fourier and Institut de Chimie Moléculaire de Grenoble (ICMG, FR-CNRS 2607), Grenoble Cedex 9, France

Background: Biodegradation of green algal cell wall requires specialized enzymatic machinery, which is not yet well characterized.
Results: Structural and biochemical characterization of a new β-glucuronyl hydrolase belonging to family GH105 active on oligo-ulvans.
Conclusion: The GH105 family encompasses enzymes cleaving both α- and β-linked glycosides.
Significance: Investigations of enzymatic degradation of marine polysaccharides reveals enzymes with unique characteristics.

Ulvans are cell wall matrix polysaccharides in green algae belonging to the genus *Ulva*. Enzymatic degradation of the polysaccharide by ulvan lyases leads to the production of oligosaccharides with an unsaturated β-glucuronyl residue located at the nonreducing end. Exploration of the genomic environment around the *Nonlabens ulvanivorans* (previously *Percicivirga ulvanivorans*) ulvan lyase revealed a gene highly similar to known unsaturated uronyl hydrolases classified in the CAZy glycoside hydrolase family 105. The gene was cloned, the protein was overexpressed in *Escherichia coli*, and enzymology experiments demonstrated its unsaturated β-glucuronyl activity. Kinetic analysis of purified oligo-ulvans incubated with the new enzyme showed that the full substrate specificity is attained by three subsites that preferentially bind anionic residues (sulfated rhamnose, glucuronic/iduronic acid). The three-dimensional crystal structure of the native enzyme reveals that a trimeric organization is required for substrate binding and recognition at the +2 binding subsite. This novel unsaturated β-glucuronyl hydrolase is part of a previously uncharacterized subgroup of GH105 members and exhibits only a very limited sequence similarity to known unsaturated β-glucuronyl sequences previously found only in family GH88. Clan-O formed by families GH88 and GH105 was singular in the fact that it covered families acting on both axial and equatorial glycosidic linkages, respectively. The overall comparison of active site structures between enzymes from these two families highlights how that within family GH105, and unlike for classical glycoside hydrolysis, the hydrolysis of vinyl ether groups from unsaturated saccharides occurs independently of the α or β configuration of the cleaved linkage.

Green algae are well known for their proliferation in eutrophicated coastal waters, resulting in so-called "green tides." One of the most spectacular green tides occurred along the shores of Qingdao just before the 2008 Olympic Games in Beijing, China (1). Due to the strong environmental impact caused by the massive proliferation of such algae, there is a need for strategies for managing and utilizing this biomass, which has high industrial potential. Current applications based on bioconversion of green algae biomass are very limited; the main present use is as a food source, consumed under the name of "sea lettuce" (*Ulva lactuca*) in Asian and Mediterranean countries. However, the cell wall matrix polysaccharides found in the genus *Ulva*, described as ulvans, have promising potential with regard to their original chemical structure and physicochemical properties (see for review, see Ref. 2). In contrast to the cell wall polysaccharides found in brown algae (*e.g.* alginates) or red algae (*e.g.* agars, carrageenans), both of which are widely used in industrial applications for their gelling and thickening properties, green algal polysaccharides have been less investigated to date.

Ulvans represent 8–30% of the dry weight of *Ulva* species. These anionic polysaccharides have a complex structure that has not yet been completely elucidated. Ulvans are mainly composed of 3-sulfated rhamnose (Rha3S),[3] glucuronic acid (GlcA), its C5 epimer iduronic acid (IduA), and in smaller amounts, xylose (Xyl) residues (3–5). Chemical and enzymatic degrada-

* This work was supported in part by the French National Research Agency (ANR; Recherche et Innovation en Biotechnologie, Grant ANR-07-RIB-019 for the Ulvoligo Project), the Pôle Mer Bretagne competitiveness cluster, the CNRS, and the Brittany Regional Council.
The atomic coordinates and structure factors (code 4CE7) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Present address: Amadeite SAS, Pôle Biotechnologique du haut du bois, F-56580 Brehan, France.
[2] To whom correspondence should be addressed: CERMAV-CNRS, Rue de la Chimie, BP53x, F-38041 Grenoble cedex 9, France. Tel.: 33-4-76-03-76-61; Fax: 33-4-76-54-72-03; E-mail: helbert@cermav.cnrs.fr.

[3] The abbreviations used are: Rha3S, 3-sulfated rhamnose; GlcA, glucuronic acid; IduA, iduronic acid; Xyl, xylose; GH, glycosyl hydrolase; GH88, GH family 88; TAIL-PCR, thermal asymmetric interlaced PCR; AD primers, arbitrary degenerate primers; Se-Met, seleno-L-methionine; Ugl, glucuronyl hydrolase.

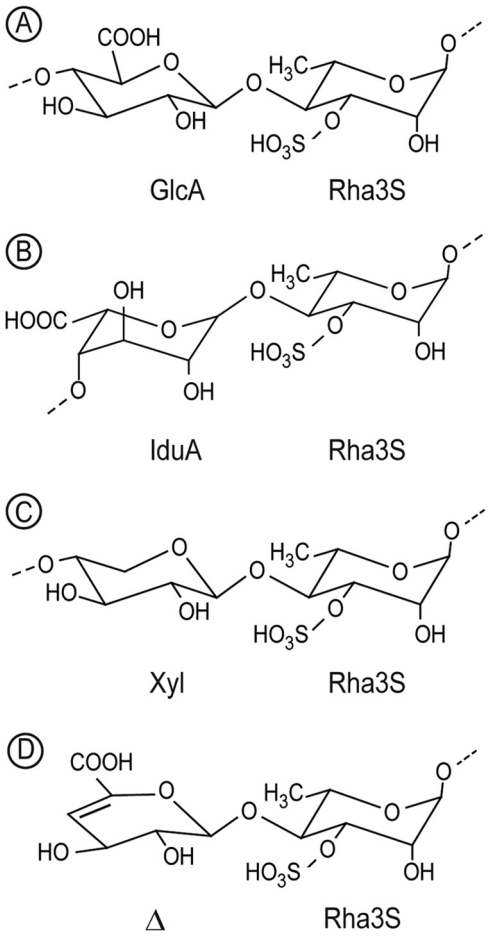

FIGURE 1. **Chemical structure of the three main disaccharide repetition moieties encountered in ulvans and the unsaturated oligosaccharide produced by lyase degradation.** *A* and *B*, ulvanobiouronic acid A and B, respectively. *C*, ulvanobiose. *D*, 4-deoxy-L-threo-hex-4-enopyranosiduronic acid (△) linked to the sulfated rhamnose occurring at the non-reducing end after cleavage of the glycosidic linkage by ulvan lyase.

tion of ulvans indicate that disaccharides composed of Rha3S, linked either to GlcA, IduA, or Xyl, giving Rha3S-GlcA (ulvanobiouronic acid A), Rha3S-IduA (ulvanobiouronic acid B), and Rha3S-Xyl, respectively, are the main repetition moieties (Fig. 1; Refs. 2, 4, and 6). Structural complexity of ulvans also increases with the possible sulfation of Xyl and with the distribution of blocks of GlcA residues. The relative proportions of these repetition moieties depend on the algal species, harvesting season, stabilization treatment, and extraction procedures (7–9).

Identification of ulvan-degrading microorganisms and of the corresponding enzymes is necessary to develop protocols for the bioconversion of green algal biomass. Microorganisms able to completely degrade green algal biomass are probably a good source of the ulvanolytic enzymes for biotechnological applications but equally integrate the disassembling tools needed to decipher ulvan structure. Ultimately, such enzymes could render possible the production and analyses of calibrated series of oligosaccharides. Accordingly, ulvan lyases have been isolated from an uncharacterized Gram-negative marine bacterium found in decomposing algae (6) and in a marine Bacteroidetes, *Nonlabens ulvanivorans*, isolated from the feces of sea hares (*Aplysia punctata*) fed with green algae (10–12). More surprisingly, ulvan lyase activity has also been isolated from a Proteobacteria species, *Ochrobactrum tritici*, found in soil (13). In all cases, the main ulvan-degrading enzymes observed thus far are ulvan lyases that catalyze the cleavage of the glycosidic bond between a sulfated rhamnose and uronic residues through a β-elimination mechanism.

The products released by polysaccharide lyases are oligosaccharides with a degree of polymerization higher than two, terminated by an unsaturated uronyl residue at the non-reducing end (14, 15). To complete the degradation of these products, microbial genomes equally encode enzymes that can further cleave of the unsaturated sugar. For instance, the end products of glycosaminoglycan lyase action are terminated by an unsaturated β-glucuronyl residue, which can be specifically digested by an unsaturated β-glucuronyl hydrolase (16, 17) found in glycoside hydrolase family 88 (GH88, CAZy database; Ref. 18). Similarly, rhamnogalacturonan lyase action yields oligosaccharides with unsaturated α-galacturonyl non-reducing ends, digested by unsaturated α-galacturonyl hydrolase (from family GH105) (19, 20). The new hydration mechanism of these enzymes was recently elegantly demonstrated by the use of synthesized substrate analogs probed by NMR, showing that the reaction directly catalyzed is the *syn* hydration of a vinyl ether to give an unstable hemiketal, and it is the collapse of this species that leads to glycosidic bond cleavage (21). Despite the fact that hydrolysis of vinyl ether groups in unsaturated saccharides occurs independently of the α- or β-configuration, to date all characterized members of either family proceed with conserved specificity with respect to the anomeric configuration of the glycosidic bond.

Our recent study describing the sequencing of the genomic environment around the ulvan lyase gene found in *N. ulvanivorans* (11) revealed a gene whose sequence is similar to uronyl hydrolases belonging to the GH105 family. The biochemical and structural characterization of this novel unsaturated uronyl hydrolase reported here reveals a first GH105 member acting specifically on the unsaturated end products of ulvan lyase action cleaving an equatorial β-glycosidic bond.

EXPERIMENTAL PROCEDURES

Sequencing of the Unsaturated β-Glucuronyl Hydrolase Gene—The sequence of the unsaturated β-glycuronyl hydrolase gene was determined using the TAIL-PCR method (22) starting from the gene of ulvan lyase. The nested specific primers used were 5′-CTATCCTTAAAAGCTGGCTCTGGAAAAGCACC-3′ (TailGH105 1R), 5′-GTACCCAACCTACACGACCATCATCATGC-3′ (TailGH105 2R), and 5′-GTATGGAAAGCACTACCACTAACCTCACCG-3′ (TailGH105 3R). Five different arbitrary degenerate primers (AD primers) were chosen from sequences found in the literature (22, 23): TGWGNAGWANCASAGA (AD1); AGWGNAGWANCAWAGG (AD2); WGTGNAGWANCANAGA (AD3); NTCGASTWTSGWGTT (AD4); NGTCGASWGANAWGAA (AD5). Primary TAIL reactions were performed in 20 μl containing 15 ng of genomic DNA, 1 μl of GoTaq PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM each of deoxyribonucleotide triphosphate (dNTP), 0.2 μM concentrations of the first specific primer, one AD primer (5 μM AD1 and

-2, 4 μM AD3, 2 μM AD4, or 3 μM AD5), and 1.25 units of GoTaq (Promega).

The conditions for the secondary TAIL reactions were identical to the first except that 1 μl of a 1:50 dilution of the primary TAIL reaction was used as a template, and the second specific primer was used in combination with the same AD primer as used in the primary TAIL reaction. For the tertiary TAIL reaction, 1 μl of a 1:50 dilution of the secondary TAIL reaction was used as template, and the third specific primer was used. The PCR programs were different for the three TAIL reactions and were based on published programs (22) but adapted to the thermocyclers available in the laboratory (11).

Heterologous Expression and Purification of the GH105 Protein—Primers were designed to amplify the GH105 gene from *N. ulvanivorans* genomic DNA and to incorporate BamHI and EcoRI restriction sites into the 5′ and 3′ ends, respectively. The forward primer sequence was GGGGGGGGATCCTGTACTGATACTGAAAAAACACCATTA, and the reverse primer CCCCCCGAATTCTTATCTCATTTTTAGAATTTCACTTCCAGC. Standard PCR conditions were used with an annealing temperature of 50 °C and 30 cycles with 10 ng of *N. ulvanivorans* genomic DNA as template. The resulting PCR products were purified, digested with appropriate restriction enzymes, and subcloned into the pFO4 expression vector, modified from pET15 (Novagen) to be compatible with the BamHI/EcoRI ligation strategy and the production of His-tagged protein, possible to purify on a Ni^{2+} chelating column. Recombinant plasmids were used to transform *Escherichia coli* strain BL21 (DE3). Transformed colonies were grown for 3 h at 37 °C in LB medium containing ampicillin and 0.5% glucose, where after an equivalent volume of cold LB medium containing 0.6% lactose, 20 mM Hepes, pH 7.0, and 1 mM isopropyl 1-thio-β-D-galactopyranoside was added, and the culture was incubated 18 h at 20 °C (24) for expression. After centrifugation, the bacterial pellet was resuspended in a buffer containing 20 mM Tris-HCl, 500 mM NaCl, and 5 mM imidazole at a pH of 7.4. The cells were lysed using a French press followed by centrifugation to remove bacterial debris. The resulting supernatant was applied to a nickel-Sepharose column charged with 100 mM $NiSO_4$ (GE Healthcare). After washing, the bound proteins were eluted with a linear gradient of imidazole ranging from 5 to 500 mM. The active fractions were pooled and injected on a Superdex 75 HiLoad prepacked column (1.6 × 60 cm; GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 8.0, with 200 mM NaCl run at 1 ml/min. During purification, the active fractions were analyzed by SDS-PAGE. Protein quantification was performed on 2-μl aliquots using a NanoDrop spectrophotometer (Thermo Scientific) using the extinction coefficient 109,780 M^{-1} cm^{-1} calculated from the amino acid sequence. Analytical gel filtration was performed by injecting 200 μl on a Superdex 200 HiLoad prepacked column (1.0 × 30 cm; GE Healthcare) run at 0.5 ml/min with the same buffer as used for the preparative gel filtration.

The seleno-L-methionine (Se-Met) labeling procedure was performed by growing recombinant *E. coli* BL21 (DE3) in 200 ml of PASM 5052 medium (25). The purification procedure was the same as the native enzyme, except that the final buffer contained 5 mM tris(2-carboxyethyl)phosphine. The Se-Met labeled *Nu*_GH105 enzyme was concentrated to 10 mg/ml.

Crystallization and Crystal Structure Determination of the GH105 Enzyme—Suitable crystallization conditions were first identified by screening 192 conditions (Qiagen JCSG+ and PACT commercial kits) using a nano-drop dispensing robot (Honeybee, Proteomics Solution) and corning 96-well crystallization plates. The best condition identified contained 25% Peg 3350 and 100 mM KNO_3. After optimization of these conditions performing a grid-screen in 24-well Linbro plates, single crystals of native and Se-Met-labeled *Nu*_GH105 were grown by mixing 2 μl of protein with 1 μl of reservoir solution containing 100 mM Tris buffer at pH 8.5, 24% (w/v) PEG MME 2000, and 0.2 M potassium nitrate and 100 mM Tris buffer at pH 8.5, 22% (w/v) PEG MME 3350, 0.2 M potassium nitrate, respectively. The hanging drops were equilibrated against 0.5 ml of reservoir solution.

Structure Determination and Refinement—Before data collection, the crystals were rapidly soaked in a cryo-buffer that was identical to the reservoir solution supplemented with 10% glycerol and subsequently frozen in a nitrogen gas stream at 100 K. Native data were collected at the European Synchrotron Radiation Facility (Grenoble, France) on beamline ID23-1 equipped with an ADSC Qunatum315 detector. The single anomalous diffraction data were collected on the beamline PROXIMA1 at the synchrotron SOLEIL (Saint-Aubin, France) at the K absorption edge of the selenium peak. The beamline is equipped with a Pilatus detector. The software program package XDS (26) was used for all data reduction and scaling. Native and Se-Met labeled *Nu*_GH105 crystals belonged to the same orthorhombic space group $P2_12_12_1$ with three molecules per asymmetric unit. All further data collection statistics are summarized in Table 1.

Phases were determined using the program PHASER (27) to find 34 selenium sites via the graphical user interface CCP4i. Initial phases, with an overall figure of merit of 0.34, were improved by noncrystallographic symmetry averaging and solvent flattening using the program PARROT (28), which provided an interpretable electron density map with an overall figure of merit of 0.872 after 15 cycles. Automatic model building with ARP/WARP (29) correctly built ~86% of the three molecules in the asymmetric unit. This model was used for molecular replacement in the native dataset at 1.9 Å resolution. The missing parts of the model were traced manually using COOT (30). The construction cycles were alternated with positional refinement using REFMAC5 (31) including all data at 1.9 Å. Solvent molecules were added with COOT, and alternative positions were added manually and refined with REFMAC5.

Spectrophotometric Enzyme Assay—A mix of unsaturated oligosaccharides obtained by degradation of ulvan by ulvan lyase was used for the protein activity assay. Activity was characterized by a decrease in the absorbance at 235 nm. Incubations were performed directly in the spectrophotometer (U-2401, Shimadzu) equipped with a temperature-controlled cuvette holder (TCC-controller 240A, Shimadzu) set at 30 °C. The standard reaction buffer was composed of 100 mM Tris-HCl, pH 7.7, 100 mM NaCl, and 100 μM ulvan oligosaccharides, and the concentration was calculated using the extinction coef-

ficient of the double bond (4800 M^{-1} cm^{-1}). A total of a 4.4 nM concentration of protein (5 μl) was added to a reaction volume of 500 μl in a quartz cuvette with a 1-cm light path unless otherwise specified. The pH optimum was determined in the pH range 6.5–9, and the standard pH buffer was replaced by 100 mM MOPS (pH 6.5, 6.9, 7.2, and 7.7) or Tris-HCl (pH 7.2, 7.7, 8.15, 8.6, and 9.0). The temperature optimum was determined in the standard reaction buffer for temperatures between 20 °C and 55 °C in 5 °C increments. The effect of salt was evaluated in the range of 0–500 mM NaCl.

Purification of Ulvan Oligosaccharides to Use as a Substrate for the GH105 Protein—Oligo-ulvans were produced using the ulvan lyase present in the culture supernatant of *N. ulvanivorans*. A vial containing 100 ml of ZoBell medium (5 g of tryptone, 1 g of yeast extract in 1 liter of 80% seawater) and 0.4% ulvan was inoculated with *N. ulvanivorans* and incubated at 20 °C at 200 rpm for 48 h before harvest by centrifugation (11). The supernatant was concentrated 8 times and buffered with 20 mM Tris-HCl, pH 8.0, 100 mM NaCl by diafiltration on a 10-kDa cutoff membrane using a 50-ml filtration cell (Amicon). This enzyme preparation (2 ml) was used to degrade 1 g of ulvan (CEVA, Pleubian, France) in 30 ml of 200 mM NaCl, 20 mM Tris-HCl, pH 9.3, at 30 °C for 24 h. The resistant fraction was removed by filtration on a 5000-Da cutoff membrane using filtration cell (Amicon). The oligosaccharide mixture was filtered (0.2 μm) and injected on three Pharmacia Superdex 30 columns (2.6 × 60 cm; GE Healthcare) in series. The elution was conducted in 50 mM $(NH_4)_2CO_3$ at 20 °C using an isocratic Gilson 306 pump working at a flow rate of 1.5 ml min^{-1}. Oligosaccharides were detected by differential refractometry (Spectra System RI-50, Thermo Separation Products), and fractions were collected using a Gilson 215 liquid handler system between 450 and 900 ml.

High Performance Anion-exchange Chromatography—The purity of the oligosaccharide fractions and the degradation kinetics of pure oligosaccharides was analyzed by high performance anion-exchange chromatography on a Dionex chromatograph ICS 3000 equipped with a 20-μl injection loop, an AS100XR automated injection system (Thermo Separation Products), and an AS11 anion exchange column (4 × 250 mm, Dionex IonPac) with an AG11 precolumn (4 × 50 mm, Dionex IonPac). The system was operated in conductivity mode using an ED40 detector (Dionex) and a Dionex ASRS ultra-4 mm suppressor with a current of 300 mA. Mobile phases were ultrapurified water and 290 mM NaOH. Elution was conducted at a flow rate of 0.5 ml min^{-1} with a GP40 gradient pump. The gradient used was 0 min, 3% B; 1.5 min, 1% B; 4.1 min, 5% B; 6.5 min, 10% B; 10.0 min, 18% B; 26 min, 22% B; 28 min, 40% B; 30 min, 100% B; 30.1 min, 3% B; 37 min, 3% B. Separation and elution of the oligosaccharides occur during the first 30 min of the gradient and is followed by a wash step and re-equilibration of the column for the next injection. The Chromeleon-peak Net software (Dionex) was used for data acquisition and transfer.

Mass Spectrometry—Protein molecular mass was determined with a Voyager DE-STR MALDI-TOF mass spectrometer (Applied Biosystems). Protein solutions (1 μl) were diluted (1/1; 1/10 and 1/100) with sinapinic acid matrix (Sigma; 10 mg ml^{-1} in 30% acetonitrile, 0.1% trifluoroacetic acid) and spotted onto the MALDI target. Spectra were acquired in positive ion linear mode under a 25-kV accelerating voltage and a mass range of 5–100 kDa. External calibration was performed using bovine serum albumin single- and double-charged ions at m/z 66,434 and 33,216 Da.

^{1}H NMR Spectroscopy—^{1}H NMR spectra were recorded at 298 K on a Bruker Avance 500 spectrometer equipped with an inverse 5-mm ^{13}C,^1H,^{15}N TCI cryoprobe. Before analysis, samples were exchanged twice in D_2O and redissolved in 99.97 atom% D_2O. Chemical shifts are expressed in ppm in reference to an external standard (trimethylsilylpropionic acid). No suppression of the HOD signal was performed.

Hierarchical Classification of Sequences—A total of 482 sequences with defined full-length catalytic modules of family GH105 were extracted from CAZy (17) on November 2013. Related sequences were grouped using a threshold of 80% sequence identity using cd-hit. To ensure that all sequences corresponding to structural or biochemical characterizations were present, a few sequences were added manually to form a set of 251 representative sequences. These sequences were aligned with Muscle 3.7, and a distance matrix was built from the multiple sequence alignment using the maximum likelihood (ML) model and the JTT substitution model (32) using an in-house modified version of Jalview. Potential subfamilies were estimated after application of the Ward hierarchical clustering method (33) to the original distance matrix. The resulting tree was analyzed under Dendroscope 3.2.9.

RESULTS

The Gene for a Divergent GH105 Glycoside Hydrolase—An ulvan lyase gene from *N. ulvanivorans* has previously been sequenced, and the heterologous expression of the corresponding protein demonstrated that it was the first representative of a novel polysaccharide lyase family (11). Here, we determined the genome sequence flanking the ulvan lyase gene using TAIL-PCR and revealed the presence of two other genes in the same reading frame. Downstream on the 3′ side, the sequence corresponds to a protein of about 70 kDa with unpredicted function. This protein was successfully overexpressed; however, we did not observe any ulvan or oligo-ulvan degradation activity (not shown). Upstream of the ulvan lyase gene, there is a gene sequence of 1130 bp (accession number JQ403607) that translated into a protein of 377 amino acids with a theoretical molecular mass of 43.75 kDa. The LipoP program predicted a clear signal peptide of 16 amino acids with cleavage between a serine in position 16 and a cysteine in position 17, suggesting that it is a lipoprotein potentially anchored on the cell wall.

Similarity searches were performed using BLASTP (34) against all non-redundant protein databases, and the 337 amino acid protein showed similarity to sequences classified in the glycoside hydrolase family GH105 found in CAZy. The highest similarities were observed with the 347-amino acid sequence of a *Formosa agariphila* KMM 3901 protein (identity 76%, accession number CDF79934) and with the 344-amino acid sequence of an *Echinicola pacifica* protein (identity 68%, accession number WP_018473370), respectively, but these proteins have not yet been biochemically characterized. On the contrary, the two characterized proteins from the GH105 family, the unsaturated

rhamnogalacturonyl hydrolases YteR and YesR (*Bacillus subtilis* subsp. *subtilis* strain 168, accession numbers, CAB14990 and CAB12519), share only 28–29% identity (over 312–335 amino acids) with the *N. ulvanivorans* sequence. These enzymes have an α-galacturonyl hydrolase activity, catalyzing the removal of the unsaturated uronic residue located at the non-reducing end of rhamnogalacturonan oligosaccharides produced by galacturonan lyases.

An alignment of the *N. ulvanivorans* sequence with the characterized GH105 sequences and those for which a three-dimensional structure is available is shown in Fig. 2. The catalytic acid/base of family GH105 enzymes that directly protonates the ring-double bond has previously been identified to be Asp-143 (numbering of YteR; Asp-148 in *Nu*_GH105 (also described as Nu_AFQ98272-4CE7) by Itoh and co-workers (20) using site-directed mutagenesis.

Biochemical Characterization of the Novel β-Glucuronyl Hydrolase—The *N. ulvanivorans* GH105 gene fused with an N-terminal histidine tag was successfully overexpressed in *E. coli* BL21 without its signal peptide. Expression levels were of ~50 mg of protein/liter of medium using standard culture and induction conditions. The protein was purified on a nickel-Sepharose chelating resin and eluted at 375 mM imidazole. The active fractions were pooled and analyzed using SDS-PAGE and MALDI-TOF MS. The most intense band migrated as expected at 40 kDa, which was measured at 43.7 kDa by mass spectrometry. However, this intense band was always associated with a second band of about 80 kDa that increased in intensity if the protein mixed with SDS sample buffer was not boiled (data not shown), suggesting incomplete denaturation of the protein. The active fractions eluted as a single peak in permeation gel chromatography on a Superdex 75 HiLoad prepacked column (1.6 × 60 cm; GE Healthcare). The active fractions were pooled and used for crystallography and biochemical analysis. Analytical gel filtration of the active fractions on a Superdex 200 (1.0 × 30 cm; GE Healthcare) separated two peaks eluting at 11.2 and 13.1 ml. Calibration of the column with known protein standards allowed determination of the peak at 13.1 ml to 128.8 kDa, which corresponds well with the expected size of a trimer. The peak at 11.2 ml is outside the range of the calibration *i.e.* larger than 200 kDa. From the genomic environment and the putative unsaturated uronyl hydrolase activity suggested by the high sequence identity with GH105 family, we suspected that oligo-ulvans were substrates. Therefore, a series of purified oligo-ulvans was incubated with the enzyme and analyzed by chromatography (Fig. 3). The unsaturated disaccharides (Δ-Rha3S) and tetrasaccharides (Δ-Rha3S-IduA-Rha3S, Δ-Rha3S-GluA-Rha3S, Δ-Rha3S-Xyl-Rha3S) were degraded into smaller products. The new oligosaccharides produced were purified, and the cleavage of the unsaturated uronic residue was confirmed by ^{1}H NMR. The signals of the ^{1}H NMR spectra of oligo-ulvans shown in Fig. 4, *A* (*bottom*) and *B* (*bottom*), were assigned based on the work of Lahaye *et al.* (6, 35). The unsaturated residue was characterized by two signals resonating at 6.04 ppm (Δ-H1) and 5.5 ppm (Δ-H4). Signals corresponding to sulfated rhamnose and uronic residues are also indicated on the spectra. After incubation with the recombinant enzyme (Fig. 4, *A* and *B*, *top*), the signals corresponding to the unsaturated residue had completely disappeared. Signals attributed to the sulfated rhamnose at the non-reducing end (Rha3Snr-H1) and to the uronic residues were slightly shifted due to a structural modification of the oligo-ulvan (Fig. 4*A*, *top*). In the case of the Rha3S-Xyl-Rha3S oligosaccharides, the removal of the unsaturated residue did not appear to modify the chemical shift of the sulfated rhamnose and the xylose residues. This suggests that the unsaturated uronic residue interacts with the internal uronic residues through ions, which cannot take place in neutral xylose.

The unsaturated glucuronyl acid residue released by the enzyme spontaneously rearranged from its cyclic form to a linear 4-deoxy-1-threo-5-hexosulose uronate. This residue thus lost its ability to absorb at 235 nm, which explains the decrease in absorbance after oligo-ulvan degradation. The optimal conditions for recombinant enzyme activity were determined using a mixture of ulvan lyase end products (11). The initial velocity as a function of temperature increased linearly from 20 to 40 °C, it remained unchanged between 40 and 45 °C, and at higher temperatures the initial velocity decreased and reached zero at 60 °C. The addition of NaCl had little effect on enzyme activity, with optimal activity obtained at 100 mM NaCl in 100 mM Tris-HCl, pH 7.7, with 80% of the activity remaining when the assay was performed in the same buffer without salt. The initial velocity as a function of pH displayed a bell shape and was centered on pH 7.7. About 60% of the activity was lost when pH was less than 6.5 or greater than 9. For comparison, the pH optimum observed with the two characterized rhamnogalacturonyl hydrolases YteR and YesR are pH 4 and 6, respectively (20).

Active Site Topology—The enzymatic degradations that were monitored by chromatography (Fig. 3) were conducted with the same concentration of oligosaccharides and with the same incubation times as used for the characterization of the recombinant protein. The three Δ-Rha3S non-reducing end-containing tetrasaccharides were completely degraded, whereas only a fraction of the corresponding disaccharides were modified, suggesting a topology that contains at least two positive binding sites and that the extra binding is important to tackle some of the substrates. The specificity of substrate recognition was further investigated by spectrophotometry, and the rate of degradation was recorded (Fig. 5). As in the chromatography analysis, the lowest rate of digestion was observed for the disaccharides (34 nM s^{-1}) followed by the Δ-Rha3S-Xyl-Rha3S tetrasaccharide (63 nM s^{-1}). Because it was not possible to obtain highly pure Δ-Rha3S-IduA-Rha3S and Δ-Rha3S-GluA-Rha3S oligo-ulvans, we analyzed the degradation of these oligosaccharides in mixtures having various compositions (Fig. 5). Despite the strong differences in tetrasaccharide composition, the observed degradation rates were very similar, varying from 107 to 114 nM s^{-1}. As a control, we assayed the degradation of rhamnogalacturonan oligosaccharides; they were not digested at all.

Crystal Structure of Nu_GH105—Crystals of native and Se-Met-labeled *Nu*_GH105 belong to space group $P2_12_12_1$ and diffract to 1.9 and 2.8 Å resolution, respectively. The Mathews coefficient V_M was calculated to be 2.7 leading to a solvent content of 54.4% with 3 molecules in the asymmetric unit. The biochemical characterization and gel-filtration experiments described above were indicative of *Nu*_GH105 being a trimer in

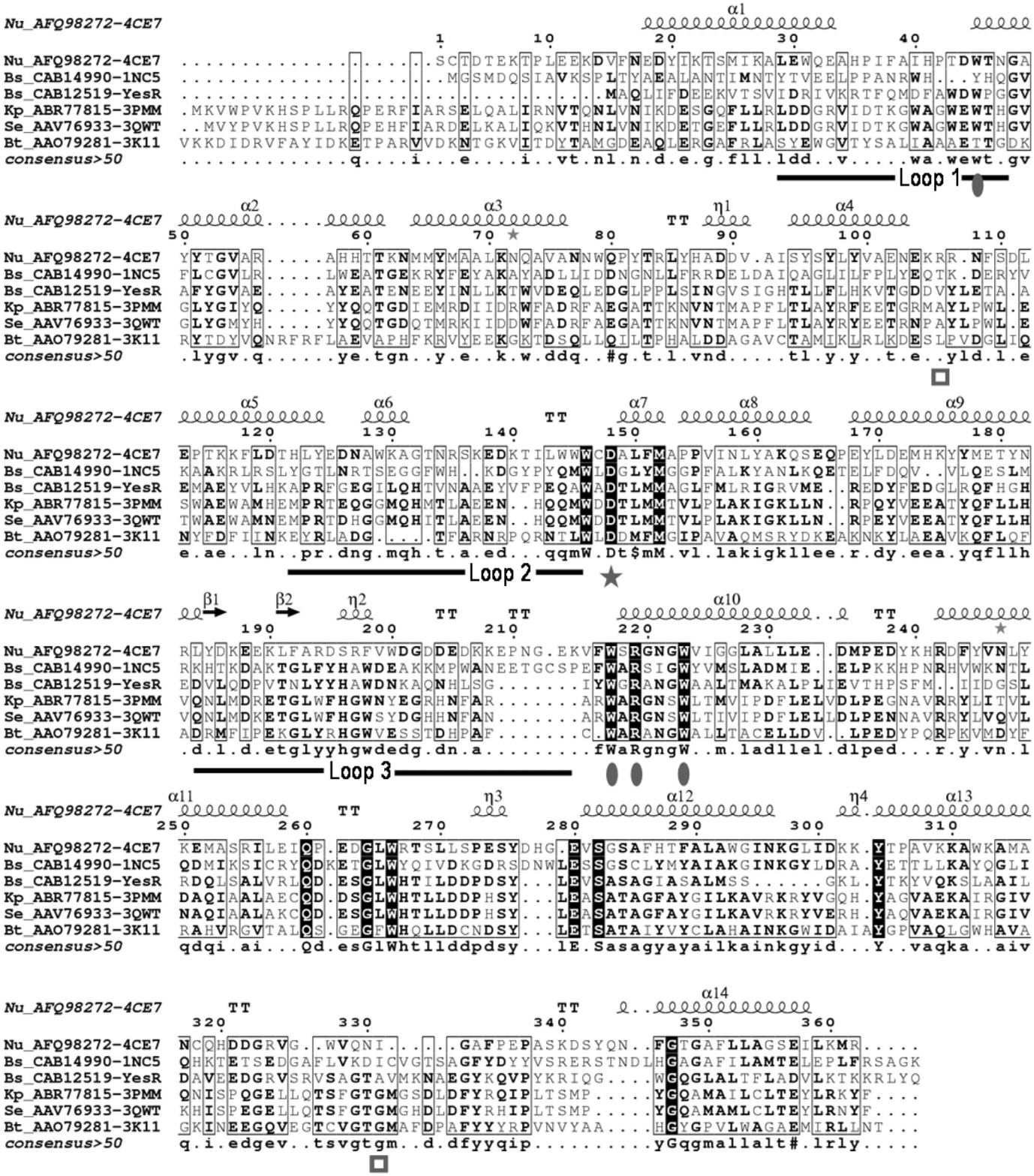

FIGURE 2. **Amino acid sequence alignment of the *N. ulvanivorans* unsaturated β-glucuronyl hydrolase sequence (*Nu*_GH105, Nu_AFQ98272–4CE7) with the two characterized proteins belonging to the GH105 family (YesR, Bs_CAB12519-YesR; YteR, Bs_CAB14990-1NC5 from *B. subtilis*) as well as the three other members of GH105 with known three-dimensional structures, Kp_ABR77815-3PMM, Se_AAV76933-3QWT, and Bt_AAO79281-3K11.** Conserved amino acids involved in the recognition of the unsaturated uronyl sugar surrounding the −1 binding site are marked by a *gray oval*; the catalytic acid/base is marked by a *star*, and two important residues in *Nu*_GH105 not conserved in other GH105 are marked by a *gray open square*. The secondary structure elements of *Nu*_GH105 are symbolized *above the sequences*, and the numbering corresponds to the mature *Nu*_GH105 enzyme. The figure was prepared using Espript (40).

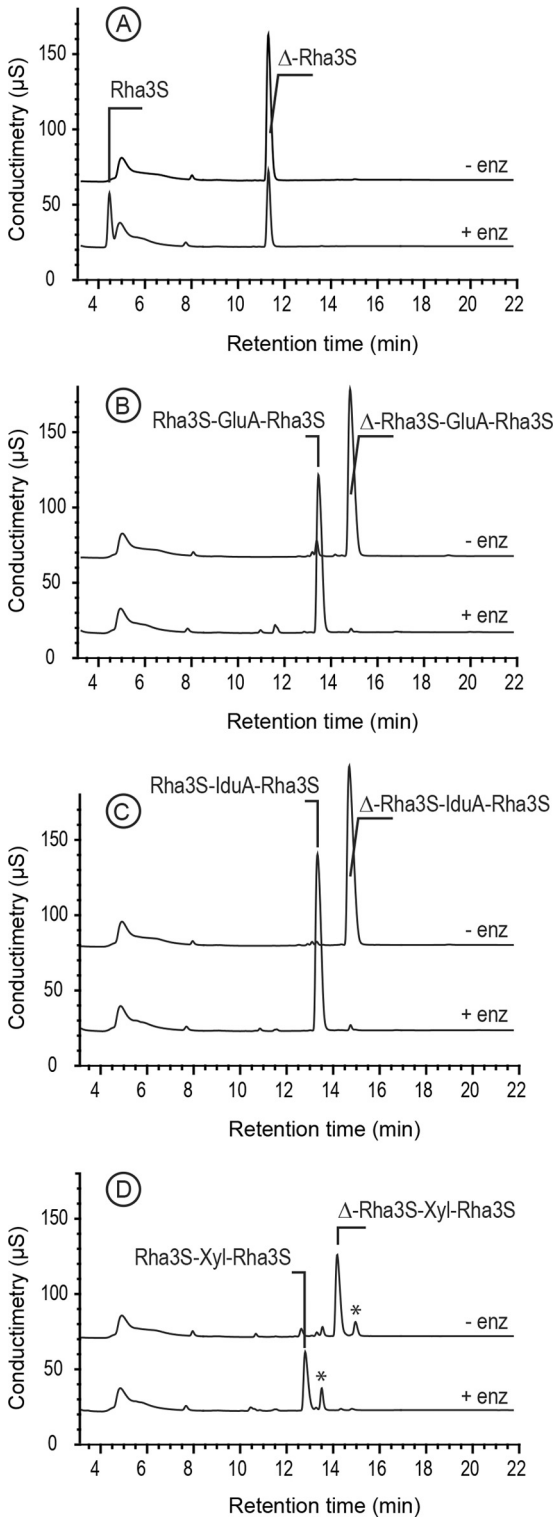

FIGURE 3. **High performance anion-exchange chromatography of purified oligosaccharides incubated with unsaturated β-glucuronyl hydrolase.** *A*, Δ-Rha3S was partially degraded by the enzyme. *B* and *C*, the tetrasaccharides Δ-Rha3S-Glc-Rha3S and Δ-Rha3S-Idu-Rha3S were completely degraded. *D*, partial degradation of the tetrasaccharide Δ-Rha3S-Xyl-Rha3S. *, uronic tetrasaccharides eluting like in *B* and *C*. −*enz*, without unsaturated β-glucuronyl hydrolase. +*enz*, with unsaturated β-glucuronyl hydrolase.

solution; we thus assume that the content of the asymmetric unit corresponds to the biological trimeric unit of *Nu*_GH105 (Fig. 6*A*). The structure was solved by the SAD method using one wavelength at the peak of the selenium adsorption edge (Table 1). The structure refinement against the native data set resulted in final R and R_{free} factors of 16.6 and 20.7%, respectively. The refined coordinates were used to perform a DALI search (36), resulting in closest matches to other GH105 enzyme members, such as 3QWT (Z-score 38.6), 2GH4 (Z-score 37.4), or 3K11 (Z-score 37.2). The root mean square deviation of the coordinates is 1.75 Å for 291 matching residues after superimposition of *Nu*_GH105 onto the structure of the characterized GH105, YteR. The enzyme adopts the $(\alpha/\alpha)_6$-barrel fold, characteristic of the GH105 or GH88 family. The active site pocket is situated at the N-terminal side of the inner six helix bundle, and many of the strictly conserved residues within family GH105 are located in this pocket. Based on our crystal structure as compared with those of YteR and YesR (19, 20), we identified several conserved amino acids, generally located in the catalytic active site pocket, surrounding the −1 binding site at the unsaturated moiety of the substrate (Fig. 7*A*). Interestingly, many of these are also conserved with enzymes belonging to family GH88, as illustrated when comparing to the unsaturated β-glucuronyl hydrolase (Ugl) from *Flavobacterium heparinum* (Fig. 7*B*). In contrast, major differences can be observed in the loops surrounding the active site pocket that contain residues forming the positive binding sites, which consequently are extremely variable throughout the three-dimensional structures of GH105 enzymes (Fig. 8).

Hierarchical clustering analysis (see tree representation shown in Fig. 9 of the GH105 representative set revealed that three distinct groups can be clearly identified in this family: (i) the first represented by the structurally and biochemically characterized Yter, (ii) a second represented by the equally characterized YesR but also with three other structural representatives without known activities, and significantly, (iii) a third set represented by the newly characterized protein *Nu*_GH105 (see below).

DISCUSSION

We have identified and overexpressed a gene with similarity to the GH105 family from the marine bacterium *N. ulvanivorans*. In agreement with its genetic location next to the recently described ulvan lyase, this novel glycosyl hydrolase is capable of degrading the unsaturated oligosaccharides produced by the ulvan lyase. The degradation kinetics suggests that the active site preferentially accommodates tetrasaccharides rather than disaccharides, which were more slowly digested. Furthermore, the presence of an uronic residue at subsite +2 was more readily accommodated than a xylose residue. Subsites −1 and +1 accommodate the unsaturated glucuronyl residue and the sulfated rhamnose, respectively, whereas the third subsite (subsite +2) can accommodate the xylose residue, but iduronic or glucuronic residues are preferred, suggesting that a specific interaction with a negatively charged carboxyl group plays a central role in the interaction. Interestingly, the enzyme was not very sensitive to the stereochemistry of the uronic units within the chain; iduronic and glucuronic units located between the two sulfated rhamnose residues were recognized equally well. Based on these kinetic experiments, we assumed that the active site must contain at least three subsites.

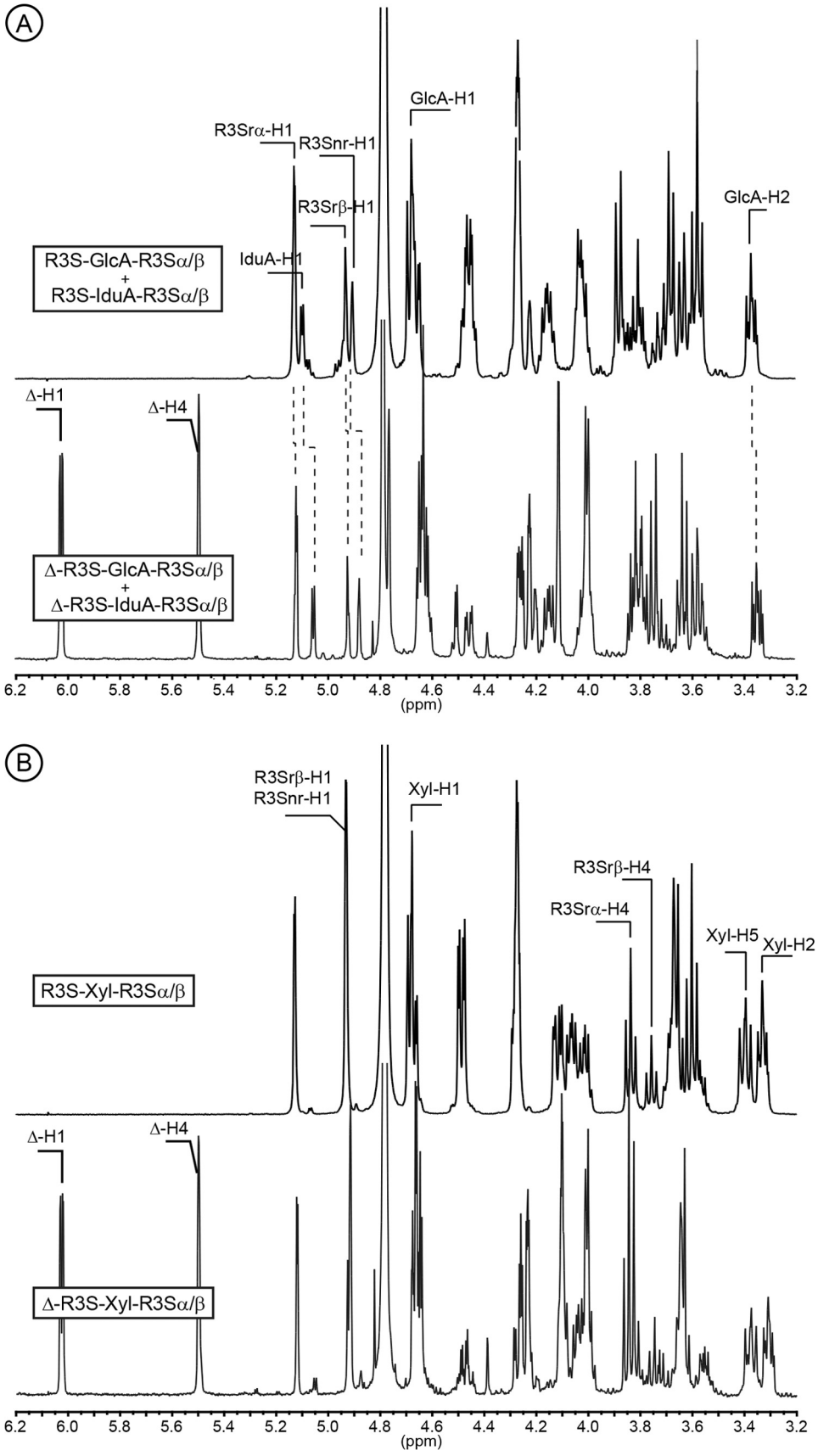

FIGURE 4. **^{1}H NMR of the purified end-products of the unsaturated β-glucuronyl hydrolase.** *A*, spectrum of a mixture of 50:50 Δ-Rha3S-Glc-Rha3S:Δ-Rha3S-Idu-Rha3S (*bottom*) and trisaccharides obtained after incubation with unsaturated β-glucuronyl hydrolase (*top*). *B*, same experiment as in *A* but with the Δ-R3S-Xyl-R3Sα/β tetrasaccharide as the starting substrate.

Regarding the crystal structure of a monomeric molecule of *Nu*_GH105, intriguingly only two binding sites can be defined within the active site pocket, namely sub-binding sites −1 and +1. Any additional sugar unit would stick out into the solvent region, outside the range of possible contacts to enzyme residues. However, when considering the trimeric arrangement of *Nu*_GH105, contacts with a neighboring symmetric molecule can stabilize additional sugar units, as illustrated in Fig. 6*B*, showing the position of a modeled substrate molecule containing a putative +2 unit in proximity of Arg-106 from a neighboring molecule. This non-conserved arginine residue of the symmetric *Nu*_GH105 molecule is ideally positioned to form a tight salt bridge with charged iduronic or glucuronic units, in agreement with the preference of charged units at the +2 binding site in the biochemical experiments. This feature appears to be unique to *Nu*_GH105 (Fig. 2), pointing toward possible substrate diversity of GH105 enzymes.

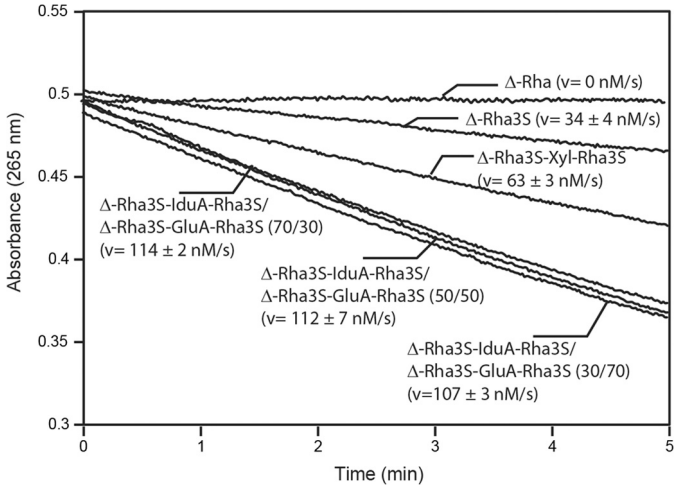

FIGURE 5. **Kinetics of degradation of purified oligosaccharides.** Structure of the unsaturated oligo-ulvans are indicated with the normalized degradation rates (s^{-1}) in *parentheses*. The activities were determined as the absorbance decrease at 235 nm determined in a quartz cuvette with a 1-cm light path at 30 °C starting with 100 μM oligosaccharide and 4.4 nM enzyme in a final volume of 500 μl.

Sequence similarity of the novel β-glucuronyl hydrolase clearly places this sequence in family GH105. However, upon alignment the sequence identities with the previously characterized members of GH105 range between 24 and 25%, reaching at most 76% with the most similar sequences of to date biochemically uncharacterized enzymes. Significantly, the sequence identity with the characterized members of GH88 ranges from 17 to 21% (over 327–362 aligned amino acid residues). Moreover, both chromatography and NMR demonstrated that the *N. ulvanivorans* enzyme is an unsaturated β-glucuronyl hydrolase able to remove the non-reducing unsaturated β-linked end Δ-Rha3S of Δ-Rha3S-IduA-Rha3S, Δ-Rha3S-GluA-Rha3S,

TABLE 1
Data collection, phasing, and refinement statistics of *Nu* G11105

Beamline	Soleil Proxima 1 07 May 2012	ESRF ID23-I 23 Fev 2011
Wavelength (Å)	0.97903	0.979
Unit cell parameters in Å	a=92.1 b=93.3 c=154.1	a=93.0 , b=93.3 , c=156.7
Space group	$P2_12_12_1$	$P2_12_12_1$
Resolution range (Å)	46.6 – 2.8 (2.92-2.78)	80.2–1.90 (2.0-1.9)[a]
No. of observations	318492 (9859)	483900 (60273)
No. of unique reflections	63791 (4590)	107955 (15114)
Completeness (%)	99.0 (96.0)	99.1 (99.1)
<I/σ(I)>	9.1 (0.9)	10.3 (1.8)
Redundancy	5.0 (2.2)	4.6 (3.9)
R_{sym}[b] ; R_{pim}[c] (%)	17.9 (65.9)	10.3 (78.9) ; 5.3 (45.1)
No of Se-sites (FOM)	33 (0.337)	
CC anomalous	29	
Refinement statistics		
Resolution range		45.2-1.9
R factor (R_{free} on 5%)		16.6 (20.7)
Overall B factor (Å^2)		28.35
No of protein atoms		8596
(mean B-factor in Å^2)		(A 27.95; B 27.35; C 28.18)
No of ions/ligand atoms		4
(mean B-factor in Å^2)		(62.0)
No of solvent atoms		723
(mean B-factor in Å^2)		(33.4)
Rms deviation in bond lengths (Å)		0.02
Rms deviation in bond angles (°)		1.894
Ramachan plot, Most favoured (%)		93.2
Ramachan plot, additional allowed (%)		6.8

[a] Values for the highest resolution shell are given in parentheses.
[b] $R_{sym} = \Sigma|I - I_{av}|/\Sigma|I|$, where the summation is over all symmetry-equivalent reflections.
[c] R_{pim} = corresponds to the multiplicity weighted R_{sym}.

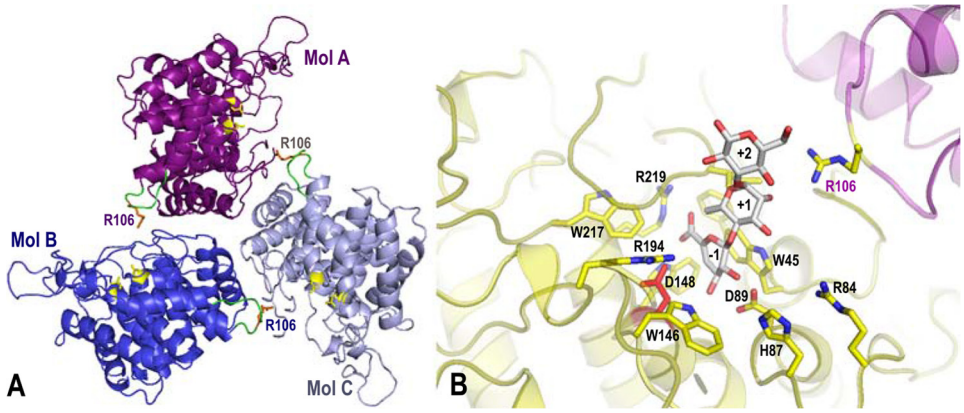

FIGURE 6. **The trimeric organization of *Nu*_GH105 provides the residues responsible for the observed substrate specificity in the +2 sub-binding site.** *A*, the crystallographic trimer organization of *Nu*_GH105. The schematic representations of the three independent molecules, arranged around the non-crystallographic 3-fold axis, are colored in *purple*, *blue*, and *gray*, respectively. The loops from residues 104–111 of each monomer are colored in *green*, and the arginine residue Arg-106 that could play a role in binding a iduronic or glucuronic acid at +2 of the *Nu*_GH105 substrate is colored in *orange*. *B*, the superimposition of *Nu*_GH105 with YteR and Ugl in complex with substrate molecules allowed the modeling of a putative substrate molecule that binds to subsites −1 to +2, highlighting the possibility that Arg-106 from a neighboring molecule could bind to a carboxylic acid function of the sixth position of a sugar unit bound to +2.

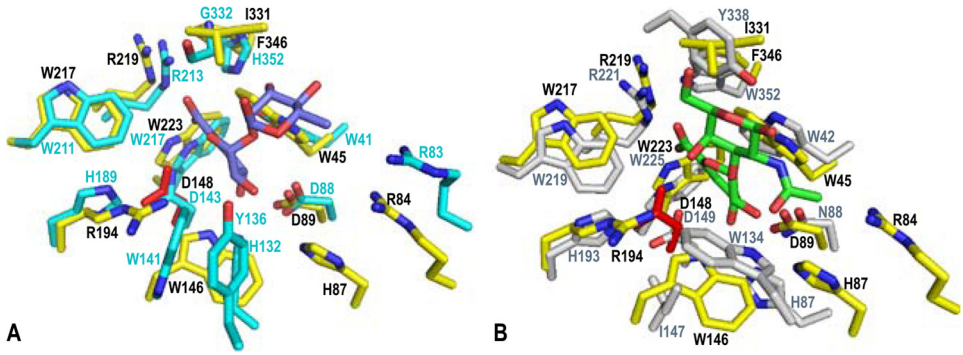

FIGURE 7. **Close-up view of the superimposed active sites of *Nu*_GH105 with the substrate complexes of YteR and Ugl.** The figure highlights that both α- and β-configured substrates could be accommodated in the active site pockets. *A*, the superimposition of *Nu*_GH105 and YteR in complex with an unsaturated α-rhamnogalacturonyl substrate (2GH4) allows identification of conserved residues and variations due to the different substrate specificities. The residues in the active site pocket of *Nu*_GH105 are colored with *yellow* carbon atoms, and the residue labels are *black*. The residues in contact with the substrate molecule in YteR are colored and labeled in *cyan*. The largest variations were observed for Ile-331 and Phe-346 (*Nu*_GH105), that are replaced by Gly-332 and His-352 in YteR, in contact with the +1 sub-binding site. Another difference was observed for Trp-146. In YteR the corresponding residue Trp-141 is in a perpendicular conformation, and the residues His-132 and Tyr-136 of YteR have no equivalents in *Nu*_GH105. *B*, the superimposition of *Nu*_GH105 and Ugl in complex with an unsaturated chondroitin disaccharide (ΔGlcA-GalNac) substrate molecule (2AHG) highlights the more divergent active site pockets, although both enzymes are active on β-configured substrate molecules. The residues in the active site pocket of *Nu*_GH105 are colored with *yellow* carbon atoms, and the residue labels are *black*. The residues in contact with the substrate molecule in Ugl are colored and labeled in *light gray*. The largest differences were again observed around the +1 binding site; Ile-331 and Phe-346 of *Nu*_GH105 were replaced by Tyr-338 and Trp-352 in Ugl, and the loop containing the residues Trp-146, His-87, and Arg-84 (*Nu*_GH105) was arranged differently in Ugl, repositioning divergent residues Ile-147, Trp-134, and His-87 on this side of the substrate binding site.

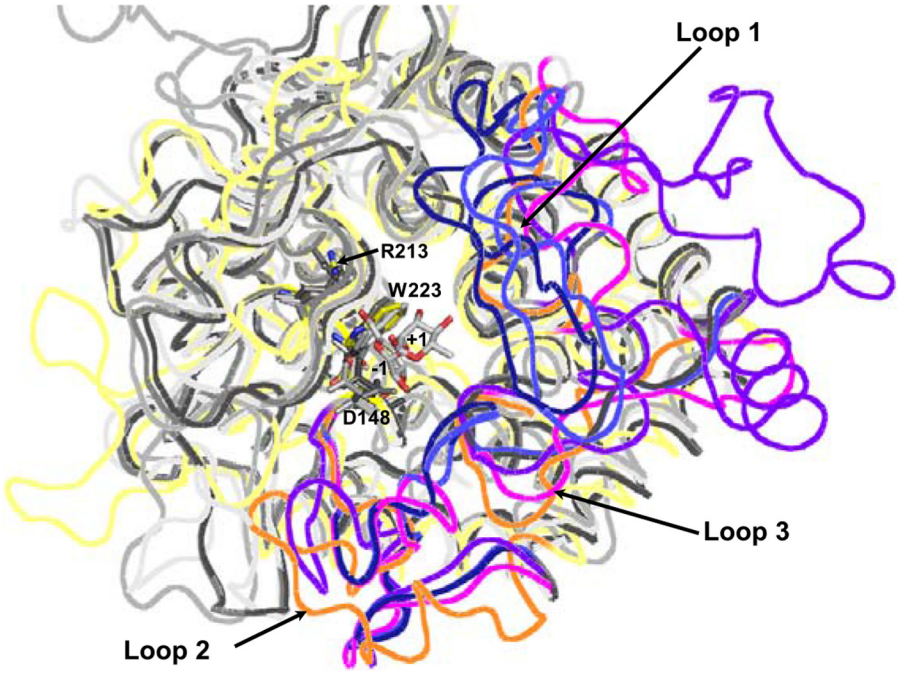

FIGURE 8. **Ribbon representation of the structural superimposition of all members of family GH105 for which crystal structures have been reported.** The backbone of the three-dimensional structure of *Nu*_GH105 is colored in *light yellow*, and the divergent loops (numbered from *1* to *3*) are highlighted in *orange*; Bs_1NC5 is colored in *light gray*, and the loops *1–3* are in *magenta*; Kp_3PMM is in *gray*, and loops *1–3* are in *bright blue*; Se_3QWT is in *medium gray*, and loops *1–3* are in *dark blue*; the backbone of Bt_3K11 is colored in *dark gray*, and loops *1–3* are in *purple-blue*.

and Δ-Rha3S-Xyl-Rha3S into a new series of fully saturated oligoulvans. Previously known unsaturated β-glucuronyl hydrolases were found only in the glycoside hydrolase GH88 family, known to hydrolyze the unsaturated disaccharides produced by lyases degrading glycosaminoglycans, such as hyaluronan and chondroitin (16). Glycosaminoglycans are linear molecules with a repeating disaccharide unit composed of a uronic acid residue linked to an amino sugar residue (glucosamine or galactosamine), which can be non-sulfated as in hyaluronan or sulfated at positions 4 and/or 6. The most striking feature of the inclusion of *Nu*_GH105 in the GH105 family is that the presently characterized enzymes of this family are known to cleave unsaturated α-anomeric galacturonyl (syn: glucuronyl) residues instead of β-anomeric uronyl residues. This constitutes the first instance of a family cleaving both equatorially and axially linked glycosidic bonds, which needs to be explained based on the singularity of the mechanism and structural information.

Both families GH105 and GH88 form clan GH-O as both structure and catalytic residues are clearly shared between these families (Fig. 8). In fact, structures from both families present an $(\alpha/\alpha)_6$ barrel fold, sharing a significant number of the amino acids next to subsite −1, including the putative cat-

FIGURE 9. **Phylogram representation of the hierarchical clustering of representative GH105 members.** Those that have been biochemically or structurally characterized (already found in Fig. 2) are labeled with their known reference GenBank™ and PDB accessions and eventually by their gene name. Three clear distinct subgroups or subfamilies can be identified with branches shown in *red*, *blue*, and *green*, respectively.

alytic aspartate amino acid, a condition that suggests the catalytic mechanism is also conserved (16–20). In both families the cleavage of the glycosidic linkage does not proceed according to classic hydrolysis mechanisms usually encountered in other glycoside hydrolases. It has been shown that cleavage starts with the addition of a proton on the carbon 4 of the unsaturated uronyl residue, which undergoes molecular rearrangements that ultimately lead to the cleavage of the glycosidic bond (37, 21). Our hypothesis is that with this mechanism, the α- or β-anomeric configurations of the unsaturated uronyl residue impose fewer constraints to the catalytic machinery and vice versa than in the case of classical retaining or inverting hydrolysis mechanisms. This is in agreement with the mechanistic study of a GH88 unsaturated β-glucuronyl hydrolase by Jongkees and Withers (21), who showed that the alternate anomer of a phenyl glycoside, having stereochemistry opposite to that of the natural substrate at carbon 1, was also accepted by Ugl as a substrate. It is, therefore, possible that the cleavage of the α- and β-anomers, both, can be observed within the same glycoside hydrolase family as it does not play a critical role for the enzymatic mechanism. This was first demonstrated by Jongkees and Withers (21) using a synthetic substrate, whereas our characterized enzyme represents the first naturally occurring example of glycosidic bonds with opposite stereochemistry being cleaved by enzymes from the same family.

The chemical structure of the rhamnogalacturonan and sulfated rhamnoglucuronan (oligo-ulvans) oligosaccharides are sufficiently similar to explain the high protein sequence identity of the corresponding uronyl degrading enzymes. The divergence observed between GH88 and GH105 protein sequences may be associated with subtle mechanistic variations in addition to different evolutionary pathways but is more likely caused by the structural and conformational differences of the substrates at the positive binding sites that are possible due to fewer constraints on the stereochemistry of the cleaved bond. This is illustrated by the sequence alignments (Fig. 2) in light of the structural superimposition of all known GH105 enzymes (Fig. 8) showing that extreme variability is observed in the loops and residues forming the positive binding sites, shaping for a large diversity of substrates even within GH105 subfamilies. We, therefore, predict that a much larger diversity of unsaturated glycuronyl-substrates than expected will be cleaved by members of this clan. Initial analysis of the tree resulting from hierarchical clustering suggests that the subgroups or subfamilies represented by the structurally and biochemically characterized Yter and by *Nu*_GH105 are more homogeneous when compared with the subgroup represented by the biochemically characterized YesR and the other remaining uncharacterized structures. The latter subgroup could, therefore, have an enlarged set of substrates or of different actions on these. Following this though, the subfamily represented by *Nu*_GH105 may only recognize very short segments of the substrate before cleavage, as suggested from the structure. In addition, it must single out that closely related sequences identified in the tree are from *Akkermansia muciniphila* and *Pedobacter heparinus*, known to be able to use mucin and heparin as carbon source, respectively (38, 39). This supports the hypothesis that very diverse substrates may be found in the corresponding subfamily.

At the moment only two other protein-coding sequences out of the >500 sequences classified in this family have been biochemically characterized, and the situation is very similar for the GH88 family, with 6 out of 306 sequences characterized. Knowledge about the substrates of additional members of the GH88 and GH105 families will be helpful to confirm the large

substrate diversity as well as the co-existence of α- and β-cleaving enzymes within the same family or even subfamily.

Furthermore, the description of this novel glycosyl hydrolase from *N. ulvanivorans* together with the recently characterized ulvan lyase leads us one step further in elucidating the degradation pathway of ulvan. These enzymatic tools can, therefore, be used to produce novel oligosaccharides or to facilitate the methanolization of green algal biomass.

Acknowledgments—We are grateful to the Biogenouest core facility for sequencing and for technical support, Stephane Cerantola at the Service Commun de Résonance Magnétique Nucléaire, Université de Bretagne Occidentale, Brest for NMR analyses, and Fanny Gaillard for mass spectrometry analyses at the Centre de Ressources Biologiques Marines, Station Biologique de Roscoff. We thank Soleil (St Auban, France) and the European Synchrotron Radiation Facilities (Grenoble, France) for according regular beam time, and we are indebted to the staff on beamline ID23-I and Pierre Legrand on Proxima 1 for technical support during data collection and treatment. The Centre de Recherches sur les Macromolécules Végétales (UPR-CNRS 5301) is affiliated with the Université Joseph Fourier and is a member of the Institut de Chimie Moléculaire de Grenoble (FR-CNRS 2607), Grenoble Cedex 9, France.

REFERENCES

1. Gao, S., Chen, X., Yi, Q., Wang, G., Pan, G., Lin, A., and Peng, G. (2010) A strategy for the proliferation of *Ulva prolifera*, main causative species of green tides, with formation of sporangia by fragmentation. *Plos ONE* **5,** 1–7
2. Lahaye, M., and Robic, A. (2007) Structure and functional properties of ulvan, a polysaccharide from green seaweeds. *Biomacromolecules* **8,** 1765–1774
3. Percival, E., and McDowell, R. H. (1967) *Chemistry and enzymology of marine algal polysaccharides*, Academic Press Inc., London
4. Quemener, B., Lahaye, M., and Bobin-Dubigeon, C. (1997) Sugar determination in ulvans by a chemical-enzymatic method coupled to high performance anion exchange chromatography. *J. Appl. Phycol.* **9,** 179–188
5. Ray, B., and Lahaye, M. (1995) Cell wall polysaccharides from the marine green alga *Ulva "rigida"* (Ulvales, Chlorophyta). Extraction and chemical composition. *Carbohydr. Res.* **274,** 313–318
6. Lahaye, M., Brunel, M., and Bonnin, E. (1997) Fine chemical structure analysis of oligosaccharides produced by an ulvan-lyase degradation of the water-soluble cell-wall polysaccharides from *Ulva* sp. (Ulvales, Chlorophyta). *Carbohydr. Res.* **304,** 325–333
7. Robic, A., Rondeau-Mouro, C., Sassi, J. F., Lerat, Y., and Lahaye, M. (2009) Structure and interactions of ulvan in the cell wall of the marine green algae *Ulva rotundata* (Ulvales, Chlorophyta). *Carbohydr. Polym.* **77,** 206–216
8. Robic, A., Sassi J. F., Dion, P., Lerat, Y., and Lahaye, M. (2009) Seasonal variability of physicochemical and rheological properties of ulvan in two Ulva species from the Brittany coast. *J. Phycol.* **45,** 962–973
9. Robic, A., Sassi, J.-F., and Lahaye, M. (2008) Impact of stabilization treatments of the green seaweed *Ulva rotundata* (Chlorophyta) on the extraction yield, the physic-chemical and rheological properties of ulvan. *Carbohydr. Polym.* **74,** 344–352
10. Barbeyron, T., Lerat, Y., Sassi, J.-F., Le Panse, S., Helbert, W., and Collén, P. N. (2011) *Persicivirga ulvanivorans* sp. *nov*., a marine member of the family Flavobacteriaceae that degrades ulvan from green algae. *Int. J. Syst. Evol. Microbiol.* **61,** 1899–1905
11. Nyvall Collén, P., Sassi, J.-F., Rogniaux, H., Marfaing, H., and Helbert, W. (2011) Ulvan lyases isolated from the Flavobacteria *Persicivirga ulvanivorans* are the first members of a new polysaccharide lyase family. *J. Biol. Chem.* **286,** 42063–42071
12. Yi, H., and Chun, J. (2012) Unification of the genera *Nonlabens*, *Persicivirga*, *Sandarakinotalea*, and *Stenothermobacter* into a single emended genus, *Nonlabens*, and description of *Nonlabens agnitus* sp. *nov*. *Syst. Appl. Microbiol.* **35,** 150–155
13. Boutachfaiti, R., Pheulpin, P., Courtois, B., and Courtois-Sambourg, J. (2009) Method for enzyme cleavage of polysaccharides derived from algae. WO 2009/016275 A2
14. Garron, M.-L., and Cygler, M. (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology* **20,** 1547–1573
15. Yip, V. L., and Withers, S. G. (2006) Breakdown of oligosaccharides by the process of elimination. *Curr. Opin. Chem. Biol.* **10,** 147–155
16. Itoh, T., Akao, S., Hashimoto, W., Mikami, B., and Murata, K. (2004) Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from *Bacillus* sp. GL1 at 1.8 Å resolution. *J. Biol. Chem.* **279,** 31804–31812
17. Maruyama, Y., Nakamichi, Y., Itoh, T., Mikami, B., Hashimoto, W., and Murata, K. (2009) Substrate specificity of streptococcal unsaturated glucuronyl hydrolases for sulfated glycosaminoglycan. *J. Biol. Chem.* **284,** 18059–18069
18. Cantarel B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active EnZymes database (CAZy). An expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238
19. Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006) Structure of unsaturated rhamnogalacturonyl hydrolase complexed with substrate. *Biochem. Biophys. Res. Commun.* **347,** 1021–1029
20. Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006) A novel glycoside hydrolase family 105. The structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *J. Mol. Biol.* **360,** 573–585
21. Jongkees, S. A., and Withers, S. G. (2011) Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* **133,** 19334–19337
22. Liu, Y. G., and Whittier, R. F. (1995) Thermal asymmetric interlaced PCR. Automatable amplification and sequencing of insert end fragments from PI and YAC Clones for chromosome walking. *Genomics* **25,** 674–681
23. Liu, Y. G., Mitsukawa, N., Oosumi, T., and Whittier, R. F. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8,** 457–463
24. Korf, U., Kohl, T., van der Zandt, H., Zahn, R., Schleeger, S., Ueberle, B., Wandschneider, S., Bechtel, S., Schnölzer, M., Ottleben, H., Wiemann, S., and Poustka, A. (2005) Large scale protein expression for proteome research. *Proteomics* **5,** 3571–3580
25. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41,** 207–234
26. Kabsch, W. (2010) *Xds*. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
27. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
28. Zhang K. Y., Cowtan, K., Main, P. (1997) Combining constraints for electron-density modification. *Methods Enzymol.* **277,** 53–64
29. Perrakis, A., Sixma, T. K., Wilson, K. S., and Lamzin, V. S. (1997) wARP. Improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 448–455
30. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
31. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
32. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8,** 275–282
33. Ward, J. H., Jr. (1963) Hierarchical grouping to optimize an objective function. *J. Am. Statist. Assoc.* **58,** 236–244
34. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008) NCBI BLAST. A better web interface. *Nucleic Acids Res.* **36,** W5–W9

35. Lahaye, M., Inizan F., and Vigouroux, J. (1998) NMR analysis of the chemical structure of ulvan and of ulvan-boron complex formation. *Carbohydr. Polym.* **36,** 239–249
36. Holm, L., and Rosenström, P. (2010) Dali server. Conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549
37. Hashimoto, W., Itoh, T., Maruyama, Y., Mikami, B., and Murata, K. (2007) Hydration of vinyl ether groups by unsaturated glycoside hydrolases and their role in bacterial pathogenesis. *Int. Microbiol.* **10,** 233–243
38. Derrien, M., Vaughan E. E., Plugge, C. M., and de Vos, W. M. (2004) *Akkermansia muciniphila* gen. *nov*., sp. *nov*., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* **54,** 1469–1476
39. Han, C., Spring, S., Lapidus, A., Del Rio, T. G., Tice, H., Copeland, A., Cheng, J. F., Lucas, S., Chen, F., Nolan, M., Bruce, D., Goodwin, L., Pitluck, S., Ivanova, N., Mavromatis, K., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. C., Saunders, E., Chertkov, O., Brettin, T., Göker, M., Rohde, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P., Detter, J. C. (2009) Complete genome sequence of *Pedobacter heparinus* type strain (HIM 762–3). *Stand. Genomic Sci.* **1,** 54–62
40. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript. Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31,** 3320–3323